a Arabidopsis. In Methods in Arabidopsis
3.

J.M., GOODMAN H.M., KOORNNEEF
MEYEROWITZ E.M., 1993. An integrated
J., 3, 745-754.

CABOCHE M., MOISAN A., JOURJON
IER D., GIRAUDAT J., GUIGLEY F.,
OKS R., GRELLET F., DELSENY M.,
LECK J., PHILIPPS G., AXELOS M.,
An inventory of 1152 expressed sequence
thaliana. Plant J., 4 (6), 1051-1061.

, SCHMIDT R., CNOPS G., DEAN C.,
ANKOFF L., SOMERVILLE C., 1991.
third of the Arabidopsis genome. Plant J.,

mapping RFLP and phenotypic markers in

OS W.D.B., HANGE B.M., GOODMAN
map of Arabidopsis thaliana. Plant Cell,

., 9, 111-127.

struction of an overlapping YAC library of
, 341-351.

Techniques et utilisations des marqueurs moléculaires
Montpellier (France), 29-31 mars 1994
Ed. INRA, Paris 1995 (Les Colloques, n°72)

Marker-assisted backcrossing: a practical example

M. RAGOT[1], M. BIASIOLLI[1], M.F. DELBUT[2], A. DELL'ORCO[1], L. MALGARINI[1], P. THEVENIN[1], J. VERNOY[2], J. VIVANT[2], R. ZIMMERMANN[1] and G. GAY[1].

1 Ciba Seeds, CH-4002 Basle, Switzerland.

2 Département d'Amélioration des Plantes, INRA-Domaine d'Epoisses, F-21110 Genlis, France

Summary

That molecular markers allow fast recovery of recurrent parent genotype in backcross programs is undisputed. Restriction Fragment Length Polymorphisms (RFLP's) were used in maize to introgress by backcross a transgene construct, containing phosphinothricin resistance and insecticidal protein genes, from a transformed parent into an elite inbred line. At each generation plants carrying the transgene construct were selected based on their phosphinothricin resistance, and further characterized with RFLP's. Both maximum recovery of recurrent parent genotype and minimum linkage drag were taken into account for marker-based selection. Embryo rescue was used to shorten generation time. Progress towards recurrent parent genotype was spectacular. Levels of recurrent parent genotype recovery which would normally be observed, in the absence of selection, in the BC_6 generation were obtained at the BC_3 generation, about one year after BC_1 seeds had been planted. Besides the evidence already provided by RFLP's, phenotypic evaluation of the backcross-derived near-isogenic lines will constitute an additional check of the completeness of the conversion.

Introduction

Backcrossing has been a common breeding practice for as long as elite germplasm has been available. It has mainly been used to introgress single Mendelian traits, such as disease resistances or quality factors, into elite germplasm (Allard 1960; Hallauer and Miranda 1981). One of the most attractive attributes of backcrossing is that it allows to perform targeted modifications without disrupting the existing overall genetic balance of the recurrent parent.

However, production of fully converted near isogenic lines through classical backcrossing procedures is a lengthy procedure, if at all possible. Theoretically, a minimum

APPENDIX 1

of seven classical backcross generations are required to recover more than 99% of recurrent parent genotype, assuming no linkage drag. The attractiveness of classical backcross procedures is therefore substantially diminished for crops, such as maize (*Zea mays* L.), where the turn-over of elite cultivars is very fast. In addition, full recovery of recurrent parent genotype is usually not achieved through classical backcrossing, which may result in deleterious agronomic effects. Murray *et al.* (1988) reported about 90% recurrent parent genotype recovery in two BC_{10}-equivalent conversions (A632Ht and A632Rp) of the maize line A632. The conversions had retained respectively 4 and 7 donor fragments in addition to the one carrying the gene of interest.

Reduction in the number of backcross generations needed to obtain fully converted individuals has been shown theoretically, or from simulations, to be achievable through the use of molecular markers (Tanksley *et al.* 1989; Hospital *et al.* 1992; Jarboe *et al.* 1994). Because they provide thorough characterization of the genetic variability at each backcross generation, markers allow to take full advantage of this variability by applying the highest possible selection intensity.

Efficiency of marker-assisted backcrossing was investigated through an experiment aimed at introgressing a single genetic factor (a transgene construct) from a donor into a recipient maize line.

Materials and methods

Plant Material

A hemizygous transgenic maize line of Lancaster origin was used as donor parent to introgress its transgene construct, through repeated backcrossing, into a recipient parent from the Stiff Stalk germplasm group. Both parents are proprietary elite lines. The transgene construct carries both a phosphinothricin resistance gene and synthetic genes encoding the entomotoxic fragment of the CryIA(b) *Bacillus thuringiensis* protein (Koziel *et al.* 1993). Transformation was achieved through microprojectile bombardment (Koziel *et al.* 1993) and resulted in a single insertion (*Bt* locus), on chromosome 1 (Figure 1).

Backcross protocol

The F1 progeny of the cross between the donor and the recipient was screened for the presence of the transgene construct by applying Basta, a phosphinothricin-based herbicide, onto each plant. Resistant individuals were then used to generate BC_1 progeny.

For each backcross generation, except the BC_4, individuals were planted in multipots and sprayed with Basta to eliminate those which did not carry the transgene construct. To avoid the stress resulting from treatment with Basta, BC_4 plants carrying the transgene construct were identified using Southern blots probed with the *pat* and *Bt* genes. Resistant plants were transplanted in an open-soil greenhouse and leaf-sampled for molecular marker

analyses. Results of marker an
flowering. A single plant was
rescued and transferred onto t
embryos first underwent a gre
culture medium, before being
average, four months.

Molecular marker analys
Restriction Fragment Le
genotypes in all four genor
chemiluminescent techniques. I
were chosen from among those
provided coverage of the entire
contained two loci tightly linke
recombination units away (Figu
BC_{n+1} generation comprised bo
or tightly linked ones, and addi
selected BC_n plant was heterozy
independent reference populati
generation.

Selection procedure
At each generation plants
recurrent-parent-genotype and
attempt to integrate both crite
missing values were not include
contributed to the selection proc
best ranking one of those for w
for the BC_3 selection) was avail

Results and discussion

Selection for the gene o
The observed segregatior
significantly different (P<0.05

Recurrent parent genoty
Statistics for the genoty
performed taking the whole g
backcross-derived plant therei

recover more than 99% of recurrent
ttractiveness of classical backcross
ops, such as maize (*Zea mays* L.).
addition, full recovery of recurrent
d backcrossing, which may result in
ported about 90% recurrent parent
(A632Ht and A632Rp) of the maize
nd 7 donor fragments in addition to

s needed to obtain fully converted
tions, to be achievable through the
al *et al.* 1992; Jarboe *et al.* 1994).
enetic variability at each backcross
variability by applying the highest

nvestigated through an experiment
ene construct) from a donor into a

origin was used as donor parent to
ckcrossing, into a recipient parent
are proprietary elite lines. The
istance gene and synthetic genes
llus thuringiensis protein (Koziel *et*
projectile bombardment (Koziel *et*
chromosome 1 (Figure 1).

the recipient was screened for the
phosphinothricin-based herbicide,
enerate BC_1 progeny.
ividuals were planted in multipots
carry the transgene construct. To
BC_4 plants carrying the transgene
th the *pat* and *Bt* genes. Resistant
eaf-sampled for molecular marker

analyses. Results of marker analyses were made available at the latest two weeks after flowering. A single plant was selected, of which all backcross-derived embryos were rescued and transferred onto tissue culture medium. Plantlets that developed from these embryos first underwent a greenhouse acclimation phase, while still growing on tissue culture medium, before being transplanted into multipots. Backcross cycles lasted, on average, four months.

Molecular marker analyses

Restriction Fragment Length Polymorphisms (RFLP's) were used to establish genotypes in all four generations. RFLP detection involved either radioactive or chemiluminescent techniques. For the BC_1 generation, 61 marker-enzyme combinations were chosen from among those revealing polymorphism between donor and recipient. They provided coverage of the entire genome, defining intervals of about 25 cM in size, and contained two loci tightly linked to the *Bt* locus, CG320 and CG415, respectively 5 and 16 recombination units away (Figure 1). For subsequent generations, markers analyzed in the BC_{n+1} generation comprised both those for which the selected BC_n plant was heterozygous, or tightly linked ones, and additional ones located in chromosomal segments for which the selected BC_n plant was heterozygous (Table 1). Marker map positions were obtained from independent reference populations and confirmed by analysis of segregation in the BC_1 generation.

Selection procedure

At each generation plants were ranked based both on the percentage of homozygous recurrent-parent-genotype and on the extent of linkage drag around the *Bt* locus, in an attempt to integrate both criteria. Plants for which two or more adjacent markers had missing values were not included in the analyses. Success or failure of the pollinations also contributed to the selection procedure. One single plant was selected at each generation: the best ranking one of those for which a backcross progeny of size 100 or more (50 or more for the BC_3 selection) was available.

Results and discussion

Selection for the gene of interest

The observed segregation ratios for phosphinothricin resistance (Table 1) were not significantly different ($P < 0.05$) from the expected 1:1, as shown by Chi-square tests.

Recurrent parent genotype recovery

Statistics for the genotyped plants are summarized in Table 1. Calculations were performed taking the whole genome into account, including the *Bt* locus. The "perfect" backcross-derived plant therefore counts one heterozygous chromosome segment, that

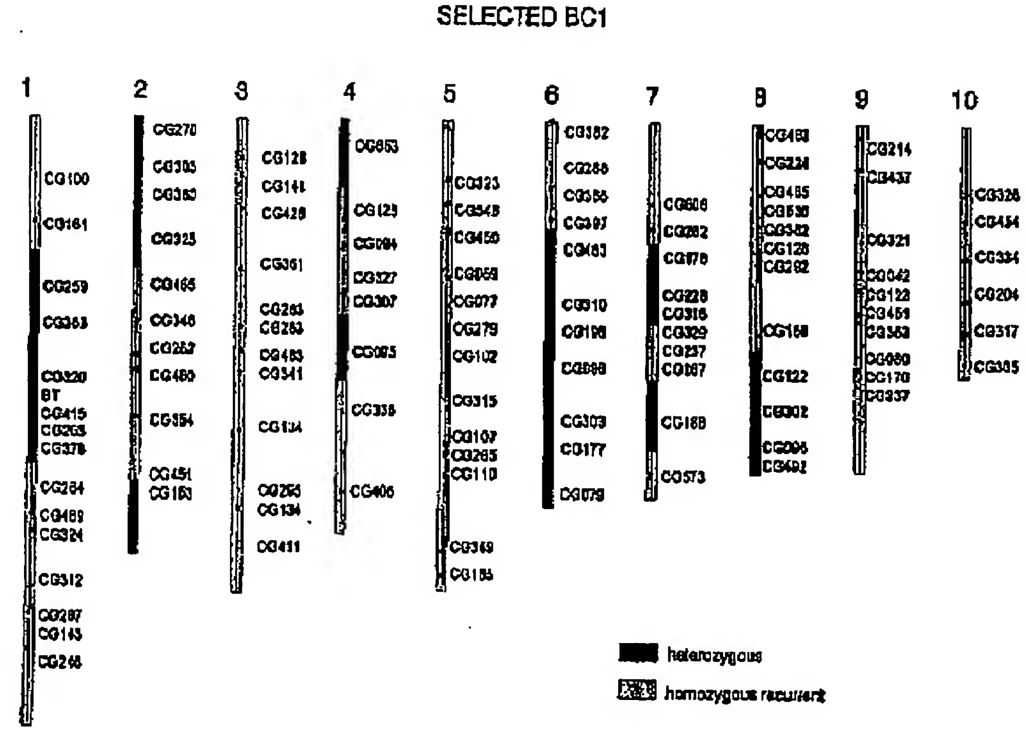
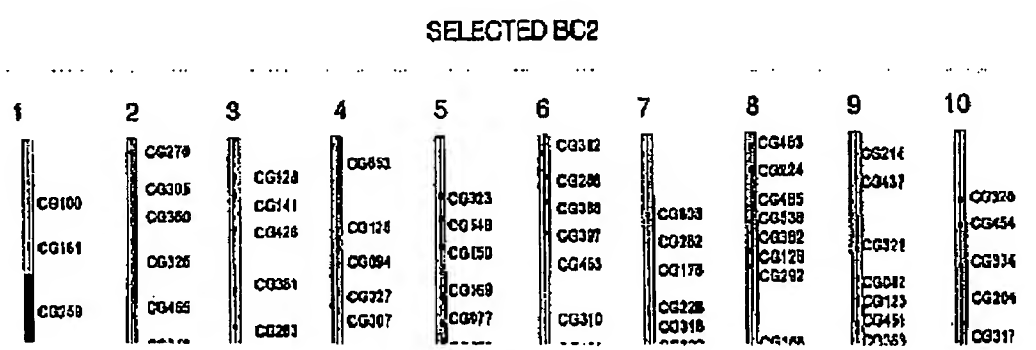


Figure 1-a: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.



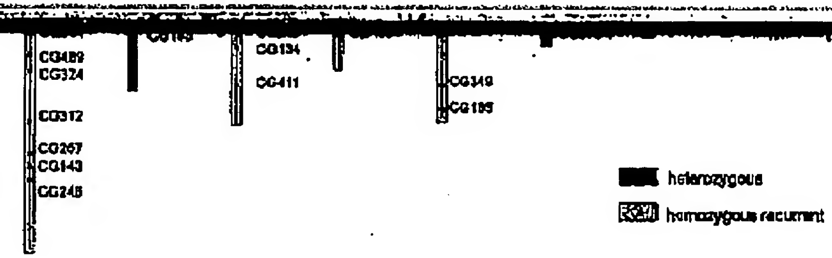


Figure 1-a: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (Bt) is located on chromosome 1.

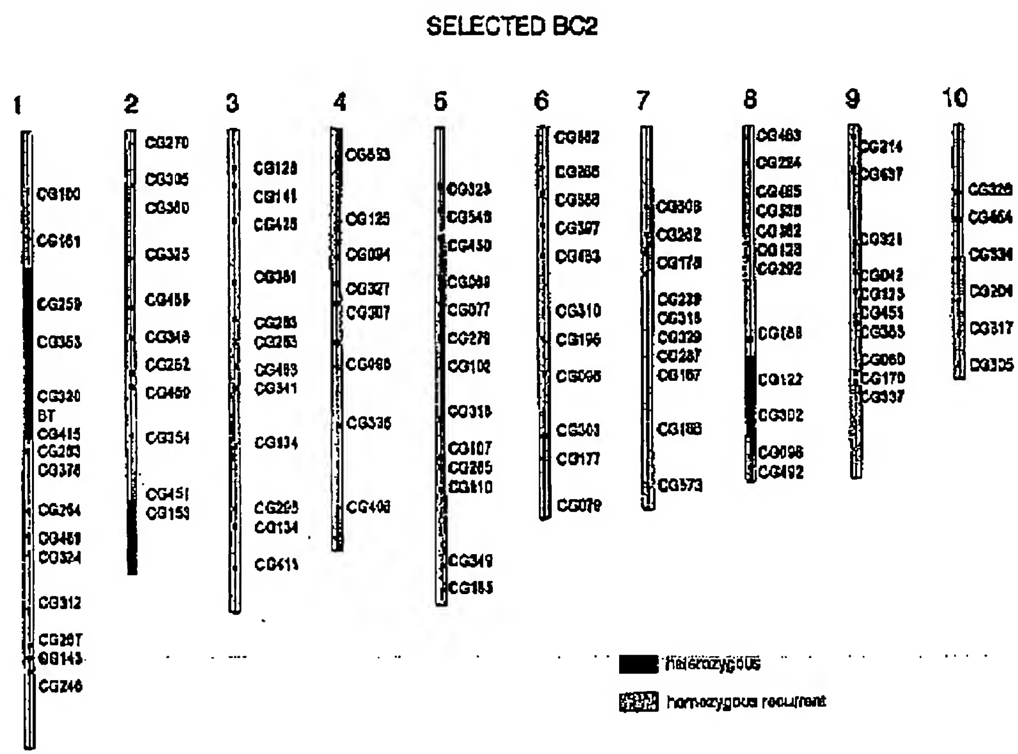


Figure 1-b: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (Bt) is located on chromosome 1.

SELECTED BC3

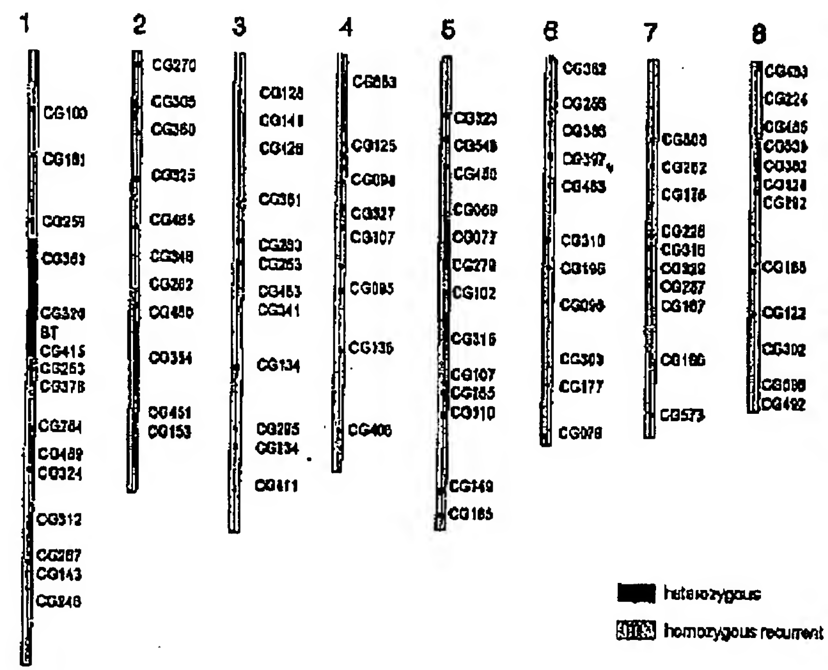
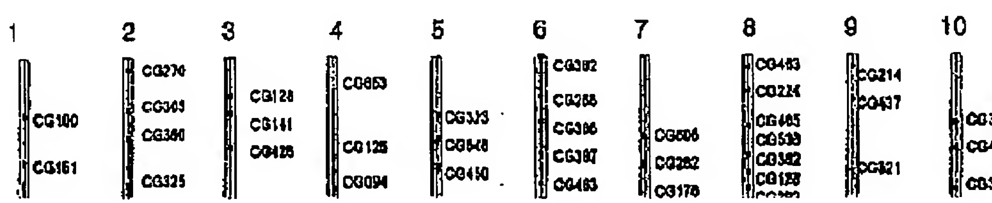


Figure 1-c: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

SELECTED BC4



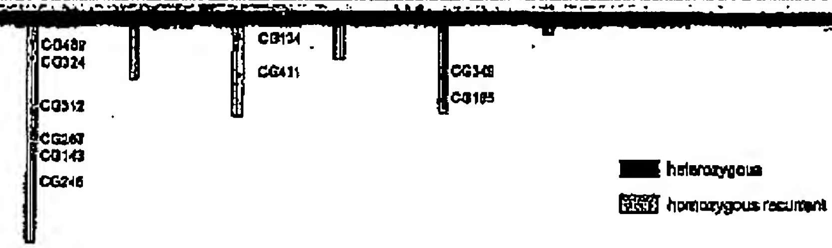


Figure 1-c: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

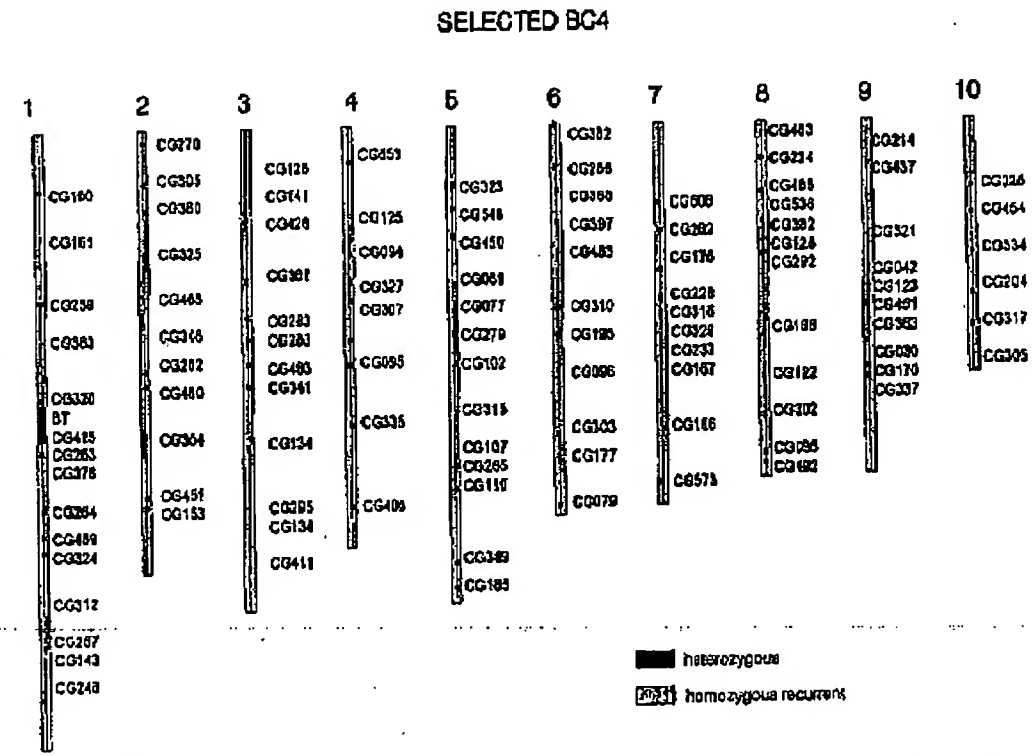


Figure 1-d: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

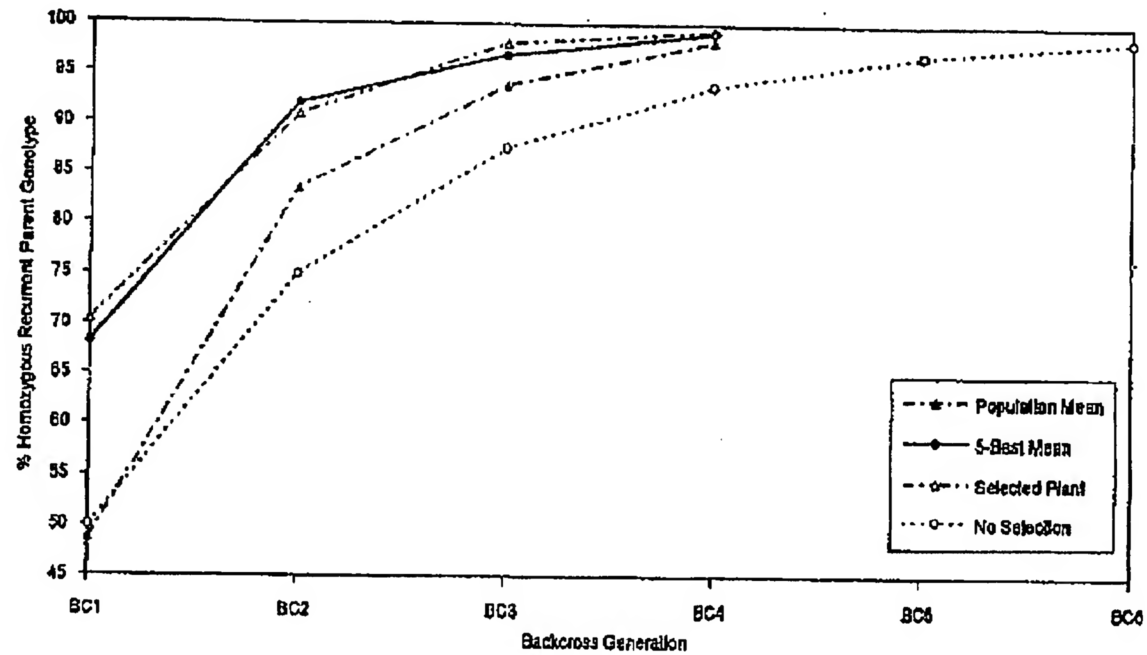


Figure 2: Recovery of recurrent parent genotype through backcrossing, with or without marker-assisted selection

Table 1: Proportion and characteristics of plants carrying the genes of interest, in the first four generations of a marker-assisted backcross program.

generation	% phosphinothricin	RFLP genotyping	nb plants	% homozygous recurrent	nb heterozygous

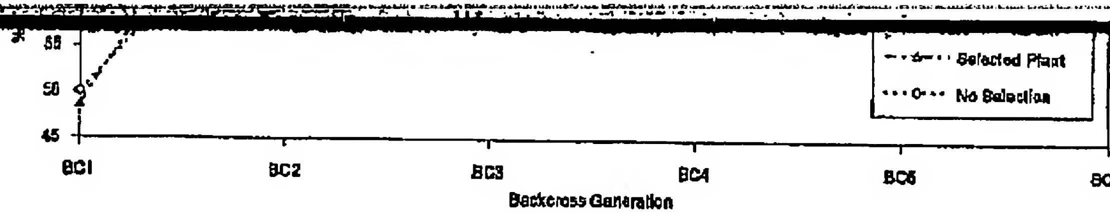


Figure 2: Recovery of recurrent parent genotype through backcrossing, with or without marker-assisted selection

Table 1: Proportion and characteristics of plants carrying the genes of interest, in the first four generations of a marker-assisted backcross program.

generation	% phosphinothricin resistant plants	RFLP genotyping			nb plants analyzed *	% homozygous recurrent parent genotype				nb heterozygous chromosome segments ***			
		nb plants	nb loci	nb datapoints		mean	std dev	5-best mean **	selected plant	mean	std dev	5-best mean **	selected plant
BC1	48.05	98	61	5868	87	48.72	10.35	88.31	70.45	11.01	2.17	7.75	6
BC2	44.65	61	22	1342	30	83.42	5.64	91.98	90.84	5.03	1.54	3.20	3
BC3	48.32	72	10	720	71	93.63	1.85	98.82	98.03	2.20	0.71	1.60	1
BC4	-	26	3	78	28	98.23	0.49	99.09	99.36	1.00	0.00	1.00	1

* Plants for which two or more adjacent markers had missing values were not included in the analyses
** Mean value of the five individuals having the five highest percentages of homozygous recurrent parent genotype.
*** Including the segment carrying the transgene construct.

comprising the *Bt* locus. It also displays 99.36% of homozygous recurrent-parent-genotype. The remaining 0.64% corresponds to the average relative length of the chromosome segment containing the *Bt* locus, which depends on the two flanking markers chosen.

The mean percentage of homozygous recurrent-parent-genotype of the BC_1 generation was slightly lower than the expected 50%. This can be explained by linkage drag around the *Bt* locus, given that this percentage was computed based only on plants selected for heterozygosity at the *Bt* locus. For all other backcross generations the mean percentage of homozygous recurrent-parent-genotype was much higher than what would have been observed, should no selection have been done (Figure 2).

The percentage of homozygous recurrent-parent-genotype of the selected plant (Table 1) and the average of the five largest values (Table 1) were always very similar to one another, and much superior to the population mean value (Figure 2). The percentage of homozygous recurrent-parent-genotype of the selected plant was found only once, in the BC_2 generation, to be smaller than the average of the five largest values. This corresponded to the only time when the selected plant was not the one with the maximum percentage of homozygous recurrent-parent-genotype. The plant had been selected because it displayed a favorable recombination on one side of the *Bt* locus (Figure 1).

The percentage of homozygous recurrent-parent-genotype of the selected BC_1 plant was almost equal to that of an unselected BC_2, that of the selected BC_2 was larger than that of an unselected BC_3, that of the selected BC_3 was barely smaller than that of an unselected BC_6, and that of the selected BC_4 was equal to that of the "perfect" backcross-derived plant, given the set of markers that was used. Such rates of recurrent parent genotype recovery are consistent with results of simulation analyses. Jarboe *et al.* (1994) who used the maize genome as a model reported that three backcross generations and 80 markers were needed to recover 99% of recurrent parent genotype.

Number of donor chromosome segments

The number of heterozygous chromosomal segments decreased from one backcross generation to the next. Plants selected at each generation were not necessarily those which had the lowest number of heterozygous chromosomal segments (Table 1). However, with the set of markers used, BC_3 and BC_4 plants were recovered which contained only one heterozygous chromosomal segment: that comprising the *Bt* locus.

Linkage drag

Linkage drag around the *Bt* locus was estimated, relative to the length of chromosome 1. Its value was found to lie between 24.0 and 48.4% for the selected BC_1 individual, between 17.6 and 34.8% for the selected BC_2, between 2.0 and 24.0% for the selected BC_3, and between 0.0 and 8.4% (respectively 0.0 and 14.5 cM) for the selected BC_4.

The two values given for each ge
correspond to extreme positions o
flanking the transgene construct locu
BC_4 is likely to be less than 1.3%
appears to be somewhat high, reflec
drag, it is much lower than what
(Stam and Zeven 1981; Tanksley et
of tomato cultivars obtained by a b
Tanksley (1989) found that the sizes
cM.

Conclusion

These results clearly demonstr
quality advantages over classical
through backcrossing. Only four ba
than a year and a half from plan
genotypically fully converted. Neve
genotype could proceed even faste
appropriate protocol and resources
allocated.

Comparison of BC_4-derived
markers and agronomic performanc
order to confirm the completeness o

References

ALLARD, R.W. (1960) Principles of plant

HALLAUER, A.R., and J.B.MIRANDA,
University Press, Ames, IA.

HOSPITAL, F., C.CHEVALET, and P.
programs. Genetics 132:1199-1210.

JARBOE, S.G., W.D.BEAVIS, and S.J.O
assisted backcross programs by compu
on the plant genome. Schergo Internat

KOZIEL, M.G., G.L.BELAND, C. BO
.DAWSON, N. DESAI, M. HILL,
McPHERSON, M.R. MEGHJI, E. MI
EVOLA (1993) Field performance of
derived from *Bacillus thuringiensis*. Bi

MURRAY, M.G., Y.MA, J.ROMERO-
fragment length polymorphisms: what

homozygous recurrent-parent-genotype.
e relative length of the chromosome
he two flanking markers chosen.
-parent-genotype of the BC_1 generation
be explained by linkage drag around the
d based only on plants selected for
ss generations the mean percentage of
higher than what would have been
e 2).
rent-genotype of the selected plant
(Table 1) were always very similar to
an value (Figure 2). The percentage of
ed plant was found only once, in the
five largest values. This corresponded
one with the maximum percentage of
d been selected because it displayed a
Figure 1).
nt-genotype of the selected BC_1 plant
f the selected BC_2 was larger than that
rely smaller than that of an unselected
at of the "perfect" backcross-derived
h rates of recurrent parent genotype
lyses. Jarboe *et al.* (1994) who used
ackcross generations and 80 markers
ype.

ments decreased from one backcross
ion were not necessarily those which
segments (Table 1). However, with
recovered which contained only one
he *Bt* locus.

relative to the length of chromosome
4% for the selected BC_1 individual,
een 2.0 and 24.0% for the selected
14.5 cM) for the selected BC_4.

The two values given for each generation are extreme values of linkage drag, which correspond to extreme positions of the crossing-overs in the marker-defined intervals flanking the transgene construct locus. Therefore the true linkage drag value of the selected BC_4 is likely to be less than 1.3% of the genome. Although this maximum value may appear to be somewhat high, reflecting the limited selection pressure put here on linkage drag, it is much lower than what would be expected from classical backcross programs (Stam and Zoven 1981; Tanksley *et al.* 1989). Practically, in a study of *Tm-2* conversions of tomato cultivars obtained by a large number of classical backcross cycles, Young and Tanksley (1989) found that the sizes of the introgressed fragments ranged between 4 and 51 cM.

Conclusion

These results clearly demonstrate that molecular markers provide important time and quality advantages over classical procedures for the production of near-isogenic lines through backcrossing. Only four backcross generations were necessary to recover, in less than a year and a half from planting of the BC_1's, individuals which appeared to be genotypically fully converted. Nevertheless, it is likely that recovery of recurrent parent genotype could proceed even faster than in the experiment described herein, should the appropriate protocol and resources (population size, number and position of markers) be allocated.

Comparison of BC_4-derived lines with the recurrent parent for both morphological markers and agronomic performance (including hybrid performance) will be performed in order to confirm the completeness of the conversion.

References

ALLARD, R.W. (1960) Principles of plant breeding. Wiley, New York, NY.

HALLAUER, A.R., and J.B.MIRANDA, Fo. (1981) Quantitative genetics in maize breeding. Iowa State University Press, Ames, IA.

HOSPITAL, F., C.CHEVALET, and P.MULSANT (1992) Using markers in gene introgression breeding programs. Genetics 132:1199-1210.

JARBOE, S.G., W.D.BEAVIS, and S.J.OPENSHAW (1994) Prediction of responses to selection in marker-assisted backcross programs by computer simulation. In: Abstracts of the second International conference on the plant genome. Scherago International Inc. 38.

KOZIEL, M.G., G.L .BELAND, C. BOWMAN, N.B. CAROZZI, R. CRENSHAW, L. CROSSLAND, J .DAWSON, N. DESAI, M. HILL, S. KADWELL, K. LAUNIS, K. LEWIS, D. MADDOX, K. McPHERSON, M.R. MEGHJI, E. MERLIN, R. RHODES, G.W. WARREN, M. WRIGHT, and S.V. EVOLA (1993) Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Bio/Technology 11:194-200.

MURRAY, M.G., Y.MA, J.ROMERO-SEVERSON, D.P.WEST, and J.H.CRAMER (1988) Restriction fragment length polymorphisms: what are they and how can breeders use them ? In: D.Wilkinson ed.,

Proceedings of the 43rd annual corn and sorghum industry research conference. American Seed Trade Association 43:72-87.

STAM, P., and C. ZEVEN (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. Euphytica 30:227-238.

TANKSLEY, S.D., N.D. YOUNG, A.H.PATERSON, and M.W.BONIERBALE (1989) RFLP mapping in plant breeding: new tools for an old science. Bio/Technology 7:257-264.

YOUNG, N.D., and S.D.TANKSLEY (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. Theor. Appl. Genet. 77:353-359.

C